

Serial No. 09/837,158
Docket No. ARC920000018US1

6

REMARKS

Claims 1-12 and 14-23 are all the claims presently pending in the application. Claims 1, 12, 14, 17-18 and 23 have been amended to more particularly define the invention. Claim 13 has been canceled.

It is noted that the claim amendments are made only for more particularly pointing out the invention, and not for distinguishing the invention over the prior art, narrowing the claims or for any statutory requirements of patentability. Further, Applicant specifically states that no amendment to any claim herein should be construed as a disclaimer of any interest in or right to an equivalent of any element or feature of the amended claim.

Applicant gratefully acknowledges the Examiner's indication that claims 12, 14 and 21 would be allowable if re-written in independent form. Applicant notes that claim 14 has been rewritten to be in independent form and is therefore in condition for immediate allowance. Applicant respectfully submits that all of the claims are allowable.

Claims 1-12 and 14-16 stand rejected under 35 U.S.C. § 101 as being allegedly directed to non-statutory subject matter.

Claims 1-6, 8, 11, 13, 17-18 and 21-23 stand rejected under 35 U.S.C. § 103(a) as being unpatentable over Allan et al. ("Topic Detection and Tracking Pilot Study Final Report", Proc. of DARPA Broadcast News Transcription and Understanding Workshop, Feb. 1998) in view of Goldszmidt et al. ("A Probabilistic Approach to Full-Text Document Clustering" 1998, Technical Report ITAD-433-MS-98-044, SRI International). Claims 7, 9-10, 15-16 and 19-20 stand rejected under 35 U.S.C. § 103(a) as being unpatentable over Allan in view of Goldszmidt, and further in view of Yang et al. ("Learning Approaches for Detecting and Tracking News Events", 1999, IEEE Intelligent Systems).

These rejections are respectfully traversed in the following discussion.

I. THE CLAIMED INVENTION

The claimed invention (e.g., as recited in claim 1) is directed to a method (e.g., a computer-implemented method) for identifying relationships between text documents and structured variables pertaining to the text documents. The method includes generating a dictionary of keywords in the text documents, forming categories of the text documents using the dictionary and an automated algorithm, counting occurrences of the structured variables, the categories, and combinations of structured variable and categories for in the text

Serial No. 09/837,158
Docket No. ARC920000018US1

7

documents. Importantly, the method also calculates probabilities of occurrences of the combinations of structured variables and categories to identify a relationship between the text documents and the structured variables.

Conventional methods of analyzing text documents cannot efficiently (e.g., automatically) identify interesting relationships between text documents (e.g., unstructured free-form text documents) and structured variables. Instead, words and phrases which frequently occur in the documents are plotted on a graph and users are required to determine for themselves whether an interesting relationship exists, which is labor intensive and time consuming (Application at page 1, line 17-page 2, line 1).

The claimed invention, on the other hand, calculates probabilities of occurrences of the combinations of structured variables and categories to identify a relationship between the text documents and the structured variables (Application at page 10, line 14-page 12, line 7). Thus, unlike conventional methods, the claimed invention can efficiently (e.g., automatically) identify interesting relationships between the structured variables and categories of text documents (Application at page 11, lines 10-11; page 23, lines 1-8).

III. THE 35 USC §101 REJECTION

The Examiner alleges that claims 1-12 and 14-16 are directed to non-statutory subject matter. Applicant notes, however, that claim 1 has been amended to recite "[a] *computer-implemented method*", which was previously recited in claim 13.

In view of the foregoing, Applicant respectfully submits that these claims are clearly directed to statutory subject matter. Therefore, the Examiner is respectfully requested to withdraw this rejection.

IV. THE ALLEGED PRIOR ART REFERENCES

A. The Allan and Goldszmidt Publications

The Examiner alleges Allan would have been combined with Goldszmidt to form the invention of claims 1-6, 8, 11, 13, 17-18 and 21-23. Applicant submits, however, that these references would not have been combined and even if combined, the combination would not teach or suggest each and every element of the claimed invention.

Allan discloses a method of event detection which is based on group average agglomerative text clustering, aiming the discovery of natural patterns of news stories over

Serial No. 09/837,158
Docket No. ARC920000018US1

8

concepts (lexicon terms) and time. The method creates a hierarchical tree of clusters, with the top layer representing a rough division into general topics and the lower layers being a finer division into narrower topics and events (Allan at section 3.2, The CMU Approach).

Goldszmidt discloses a probabilistic approach to full-text document clustering which includes scoring document similarity based on probabilistic considerations. Similarity is scored according to the expectation of the same words appearing in two documents. The score enables the investigation of different smoothing methods for estimating the probability of a word appearing in a document, for purposes of clustering (Goldszmidt at Abstract).

Applicant respectfully submits that these references would not have been combined as alleged by the Examiner. Indeed, Allan is directed to a method of event detection which is based on group average agglomerative text clustering, whereas Goldszmidt is directed to a method which estimates the probability of a word appearing in a document, for purposes of document clustering. Thus, these references are completely unrelated, and no person of ordinary skill in the art would have considered combining these disparate references, absent impermissible hindsight.

Further, these references clearly do not teach or suggest their combination. Therefore, Applicant respectfully submits that one of ordinary skill in the art would not have been so motivated to combine the references as alleged by the Examiner. Therefore, the Examiner has failed to make a prima facie case of obviousness.

Moreover, neither Allan, nor Goldszmidt, nor any alleged combination teaches or suggests a method for identifying relationships between text documents and structured variables pertaining to the text documents, which includes "calculating probabilities of occurrences of said combinations of structured variables and categories to identify a relationship between said text documents and said structured variables", as recited, for example, in claims 1 and 23, and similarly recited in claim 17.

As noted above, unlike conventional methods of analyzing text documents in which words and phrases which frequently occur in the documents are plotted on a graph and users are required to determine for themselves whether an interesting relationship exists, the claimed invention, on the other hand, calculates probabilities of occurrences of the combinations of structured variables and categories to identify a relationship between the text documents and the structured variables (Application at page 10, line 14-page 12, line 7). Thus, unlike conventional methods, the claimed invention can efficiently (e.g., automatically)

Serial No. 09/837,158
Docket No. ARC920000018US1

9

identify interesting relationships between the structured variables and categories of text documents (Application at page 11, lines 10-11; page 23, lines 1-8).

Clearly, these novel features are not taught or suggested by the cited references or their combination. Indeed, the Examiner expressly concedes that Allan does not teach or suggest this feature on page 4 of the Office Action. The Examiner alleges that the feature is taught by Goldszmidt. However, the Examiner is incorrect.

Specifically, the Examiner claims that Goldszmidt teaches a document similarity metric that is based on probability. However, even assuming (arguendo) that Goldszmidt teaches a document similarity metric that is based on probability, his document similarity metric does not teach or suggest the claimed invention.

Indeed, the claimed invention is not necessarily intended to measure similarity between documents. Instead, an exemplary aspect of the claimed invention may involve looking for a correlation between some event in the document text (e.g., the occurrence of a text category or word) and some other structured variable (e.g., date of the document, author's country of origin, etc.).

The probability measure in the claimed invention is simply one way that the claimed invention may determine whether any relationship (e.g., correlation) might exist between two supposedly independent variables (text and structure) and if so, what that correspondence is (e.g. certain words or phrases used more often at certain times than at other times).

Thus, Goldszmidt does not teach or suggest a method of identifying relationships between text documents and structured variables as in the claimed invention. Specifically, Goldszmidt does not teach or suggest a method which calculates probabilities of occurrences of the combinations of structured variables and categories to identify a relationship between the text documents and the structured variables.

Therefore, even assuming (arguendo) that these references would have been combined, and even assuming (arguendo) that the Examiner's allegations regarding the teachings of Goldszmidt are correct, the combination of Allan and Goldszmidt clearly does not teach or suggest the claimed invention.

Therefore, Applicant submits that these references would not have been combined and even if combined, the combination would not teach or suggest each and every element of the claimed invention. Therefore, the Examiner is respectfully requested to withdraw this rejection.

Serial No. 09/837,158
Docket No. ARC920000018US1

10

B. The Yang Publication

The Examiner alleges Allan would have been combined with Goldszmidt, and that the alleged Allan/Goldszmidt would have been further combined with Yang to form the invention of claims 7, 9-10, 15-16 and 19-20. Applicant submits, however, that these references would not have been combined and even if combined, the combination would not teach or suggest each and every element of the claimed invention.

Yang discloses a method for detecting and tracking news events. Specifically, the method allegedly extends existing supervised-learning and unsupervised-clustering algorithms to allow document classification based on the information content and temporal aspects of news events (Yang at page 32, middle and right columns).

Applicant respectfully submits that these references would not have been combined as alleged by the Examiner. Indeed, contrary to Allan and Goldszmidt, Yang is directed to a method of event detection and tracking. Thus, these references are unrelated, and no person of ordinary skill in the art would have considered combining these disparate references, absent impermissible hindsight.

Further, these references clearly do not teach or suggest their combination. Therefore, Applicant respectfully submits that one of ordinary skill in the art would not have been so motivated to combine the references as alleged by the Examiner. Therefore, the Examiner has failed to make a prima facie case of obviousness.

Moreover, neither Allan, nor Goldszmidt, nor Yang, nor any alleged combination teaches or suggests a method for identifying relationships between text documents and structured variables pertaining to the text documents, which includes "calculating probabilities of occurrences of said combinations of structured variables and categories to identify a relationship between said text documents and said structured variables", as recited, for example, in claims 1 and 23, and similarly recited in claim 17.

As noted above, unlike conventional methods of analyzing text documents in which words and phrases which frequently occur in the documents are plotted on a graph and users are required to determine for themselves whether an interesting relationship exists, the claimed invention can efficiently (e.g., automatically) identify interesting relationships between the structured variables and categories of text documents (Application at page 11, lines 10-11; page 23, lines 1-8).

Serial No. 09/837,158
Docket No. ARC920000018US1

11

Clearly, these novel features are not taught or suggested by Yang. Indeed, Applicant would point out that the Examiner is not attempting to rely on Yang as allegedly teaching this feature.

In fact, Yang is clearly unrelated to the claimed invention. Indeed, Yang states that the tasks of his method are "1. Segmenting speech-recognized TV and radio broadcasts into news stories, 2. Detecting events from unsegmented or segmented news streams, and 3. Tracking stories for particular events based on user-identified sample stories" (Yang at page 32, right hand column).

Thus, clearly, Yang does not identify as a task of his method, identifying a relationship between the text documents and the structured variables. Certainly, Yang does not teach or suggest identifying such a relationship by calculating probabilities of occurrences of the combinations of structured variables and categories. Thus, Yang clearly does not make up for the deficiencies of the alleged Allan and Goldszmidt combination.

Therefore, Applicant submits that these references would not have been combined and even if combined, the combination would not teach or suggest each and every element of the claimed invention. Therefore, the Examiner is respectfully requested to withdraw this rejection.

III. FORMAL MATTERS AND CONCLUSION

In view of the foregoing, Applicant submits that claims 1-12 and 14-23, all the claims presently pending in the application, are patentably distinct over the prior art of record and are in condition for allowance. The Examiner is respectfully requested to pass the above application to issue at the earliest possible time.

Should the Examiner find the application to be other than in condition for allowance, the Examiner is requested to contact the undersigned at the local telephone number listed below to discuss any other changes deemed necessary in a telephonic or personal interview.

Serial No. 09/837,158
Docket No. ARC920000018US1

12

The Commissioner is hereby authorized to charge any deficiency in fees or to credit any overpayment in fees to Assignee's Deposit Account No. 09-0441.

Respectfully Submitted,

Date: 4/24/05

Phillip E. Miller, Esq.
Registration No. 46,060

McGinn & Gibb, PLLC
8321 Old Courthouse Road, Suite 200
Vienna, VA 22182-3817
(703) 761-4100
Customer No. 21254

CERTIFICATE OF FACSIMILE TRANSMISSION

I hereby certify that the foregoing was filed by facsimile with the United States Patent and Trademark Office, Examiner James Blackwell, Group Art Unit #2176 at fax number (703) 872-9306 this 26th day of April, 2005.



Phillip E. Miller
Reg. No. 46,060